

# Quantum-Space Attacks

Ran Gelles

Tal Mor

Technion - Israel Institute of Technology  
Computer Science Department  
{gelles, talmo}@cs.technion.ac.il

February 5, 2008

## Abstract

Theoretical quantum key distribution (QKD) protocols commonly rely on the use of qubits (quantum bits). In reality, however, due to practical limitations, the legitimate users are forced to employ a larger quantum (Hilbert) space, say a quhexit (quantum six-dimensional) space, or even a much larger quantum Hilbert space. Various specific attacks exploit of these limitations. Although security can still be proved in some very special cases, a general framework that considers such realistic QKD protocols, *as well as* attacks on such protocols, is still missing.

We describe a general method of attacking realistic QKD protocols, which we call the ‘quantum-space attack’. The description is based on assessing the enlarged quantum space actually used by a protocol, the ‘quantum space of the protocol’. We demonstrate these new methods by classifying various (known) recent attacks against several QKD schemes, and by analyzing a novel attack on interferometry-based QKD.

## 1 Introduction

Quantum cryptography has brought us new ways of exchanging a secret key between two users (known as Alice and Bob). The security of such Quantum Key Distribution (QKD) methods is based on a very basic rule of nature and quantum mechanics—the “no-cloning” principle. The first QKD protocol was suggested in a seminal paper by Bennett and Brassard [5] in 1984, and is now known as BB84. During recent years many security analyses were published [46, 35, 6, 42, 7, 22] which proved the information-theoretical security of the BB84 scheme against the most general attack by an unlimited adversary (known as Eve), who has full control over the quantum channel<sup>1</sup>. Those security proofs are limited as they always consider a theoretical QKD that uses perfect qubits. Although these security proofs do take errors into account, and the protocols use error correction and privacy amplification (to compensate for these errors and for reducing any partial knowledge that Eve might have), in general, they avoid security issues that arise from the implementation of qubits in the *real world*.

A pivotal paper by Brassard, Lütkenhaus, Mor, and Sanders [12, 13] presented the “Photon Number Splitting (PNS) attack” and exposed a security flaw in experimental and practical QKD: One must take into account the fact that Alice does not generate perfect qubits (2 basis-states of a single photon), but, instead, generates states that reside in an enlarged Hilbert space (we call it “quantum space” here), of six dimensions. The reason for that discrepancy in the size of the used quantum space is that each electromagnetic pulse that Alice generates contains (in addition to the two dimensions spanned by the single-photon states) also a

<sup>1</sup>All QKD protocols assume that Alice and Bob also use an insecure, yet unjammable, classical channel.

vacuum state and three 2-photon states, and these are extremely useful to the eavesdropper. That paper proved that, in contrast to what was assumed in previous papers, Eve can make use of the enlarged space, and get a lot of information on the secret key, sometimes even full information, without inducing any noise. Many attacks on the practical protocols then followed (e.g., [25, 26, 24, 36, 34, 21]), based on extensions of the quantum spaces, exploring various additional security flaws; other papers [25, 40, 45] suggested possible ways to overcome such attacks. On the one hand, several security proofs, considering specific imperfections, were given for the BB84 protocol [24, 27]. Yet on the other hand, it is generally impossible now to prove the security of a practical protocol, since *a general framework* that considers such realistic QKD protocols, *and* the possible attacks on such protocols, is still missing.

We show that the PNS attack, and actually all attacks directed at the channel, are various special cases of a general attack that we define here, the *Quantum-Space Attack* (QSA). The QSA generalizes existing attacks and also offers novel attacks. The QSA is based on the fact that the “qubits” manipulated in the QKD protocol actually reside in a larger Hilbert space, and this enlarged space *can be assessed*. Although this enlarged space is not fully accessible to the legitimate users, they can still analyze it, and learn what a fully powerful eavesdropper can do. We believe that this assessment of the enlarged “quantum space of the protocol” is a vital step on the way to proving or disproving the unconditional security of practical QKD schemes. We focus on schemes in which the quantum communication is uni-directional, namely, from Alice’s laboratory (lab) to Bob’s lab. We consider an adversary that can attack all the quantum states that come out of Alice’s lab, and all the quantum states that go into Bob’s lab.

The paper is organized as follows: Definitions of the quantum spaces involved in the realization of a protocol, and of the “quantum space of the protocol”, are presented and discussed in Section 2. The “quantum-space attack” is defined and discussed in Section 3. Using the general framework when the information carriers are photons is discussed in Section 4. Next, in Section 5 we show that the best known attacks on practical QKD are special cases of the QSA. Section 6 demonstrates and analyzes a novel QSA on an interferometric implementation of the BB84 and the six-state QKD protocols. Last, we discuss a few subtleties and open problems for future research in Section 7.

We would like to emphasize that our (crypt)analysis presents the difficulty of proving unconditional security for practical QKD setups, yet also provides an important (probably even vital) step in that direction.

## 2 The Quantum Space of the Protocol

The Quantum Space Attack (QSA) is the most general attack on the quantum channel that connects Alice to Bob. It can be applied to any realistic QKD protocol, yet here we focus on uni-directional schemes and on implementations of the BB84 protocol and the six-state protocol. We need to have a proper model of the protocol in order to understand the Hilbert space that an unlimited Eve can attack. This space has never been analyzed before except for specific cases. Our main finding is a proper description of this space, which allows, for the first time, defining the most general eavesdropping attack on the channel. We start with a model of a practical “qubit”, continue with understanding the spaces used by Alice and Bob, and end by defining the relevant space, the *Quantum Space of the Protocol* (QSoP), used by Eve to attack the protocol. The attacks on the QSoP are what we call *Quantum-Space Attacks*.

### 2.1 Alice’s realistic space

In most QKD protocols, Alice sends Bob qubits, namely, states of 2 dimensional quantum spaces ( $H_2$ ). A realistic view should take into account any deviation from theory, caused by Alice’s equipment. For example, Alice might encode the qubit via a polarized photon:  $|0_z\rangle$  via a photon polarized horizontally, and

$|1_z\rangle$  polarized vertically. This can be written using Fock notation<sup>2</sup> as  $|n_h, n_v\rangle^F$  where  $n_h$  ( $n_v$ ) represents the number of horizontal (vertical) photons; then  $|0_z\rangle \equiv |1, 0\rangle^F$  and  $|1_z\rangle \equiv |0, 1\rangle^F$ . When Alice's photon is lost within her equipment (or during the transmission), Bob gets the state  $|0, 0\rangle^F$ , so that Alice's realistic space becomes  $H_3$ . Alice might send multiple photons and then  $H^A$  is of higher dimension, see Section 4.2.

**Definition 1.** Alice's realistic space,  $H^A$ , is the minimal space containing the actual quantum states sent by Alice to Bob during the QKD protocol.

In the BB84 protocol, Alice sends qubits in two<sup>3</sup> fixed conjugate bases. Theoretically, Alice randomly chooses a basis and a bit value and sends the chosen bit encoded in the appropriate chosen basis as a state in  $H_2$  (e.g.  $|0_z\rangle, |1_z\rangle, |0_x\rangle = (|0_z\rangle + |1_z\rangle)/\sqrt{2}$ , and  $|1_x\rangle = (|0_z\rangle - |1_z\rangle)/\sqrt{2}$ ). To a better approximation, the states sent by Alice are four different states  $|\psi_i\rangle_A$  ( $i = 1, 2, 3, 4$ ) in her realistic space  $H^A$ , spanned by these four states. This space  $H^A$  is of dimension  $|H^A|$ , commonly between 2 and 4, depending on the specific implementation. As practical instruments often diverge from theory, Alice might send quite different states. As an extreme example, see the *tagging attack* (Section 5.2), which is based on the fact that Alice's space could contain more than just these four theoretical states, so that  $|H^A| > 4$  is possible.

## 2.2 Extension of Alice's space

Bob commonly receives one of several possible states  $|\psi_i\rangle_A$  sent by Alice, and measures it. The most general measurement Bob can perform is to add an ancilla, perform a unitary transformation on the joint system, perform a complete measurement, and potentially “forget”<sup>4</sup> some of the outcomes<sup>5</sup>. However, once Alice's space is larger than  $H_2$ , the extra dimensions provided by Alice could be used by Bob for his measurement, *instead of* adding an ancilla. Interestingly, by his measurement Bob might be *extending* the space vulnerable to Eve's attack well beyond  $H^A$ . This is possible since in many cases the realistic space,  $H^A$ , is embedded inside a larger space  $M$ .

**Definition 2.** The space  $M$  is the space in which  $H^A$  is embedded,  $H^A \subseteq M$ . The space  $M$  is the actual space available for Alice and an Eavesdropper.

Due to the presence of an eavesdropper, Bob's choice whether to add an ancilla or to use the extended space  $M$  is vital for security analysis. In the first case the ancilla is added by Bob, inside his lab, while in the second it is controlled by Alice, transferred through the quantum channel and exposed to Eve's deeds. Eve might attack the extended space  $M$ , and thus have a different effect on Bob, considering his measurement method.

For example, suppose Alice sends two non-orthogonal states of a qubit,  $\theta_0 = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$  and  $\theta_1 = \begin{pmatrix} \cos \theta \\ -\sin \theta \end{pmatrix}$ , with a fixed and known angle  $0 \geq \theta \geq 45^\circ$ . Bob would like to distinguish between them, while allowing inconclusive results sometimes, but no errors [38]. Bob can add the ancilla  $|0\rangle_{Anc} \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{Anc}$  and perform the following transformation  $\mathcal{U}$ :

$$\begin{aligned} |0\rangle_{Anc} \otimes \begin{pmatrix} \cos \theta \\ \pm \sin \theta \end{pmatrix} &= \begin{pmatrix} \cos \theta \\ \pm \sin \theta \\ 0 \\ 0 \end{pmatrix} \xrightarrow{\mathcal{U}} \begin{pmatrix} \sin \theta \\ \pm \sin \theta \\ \sqrt{\cos 2\theta} \\ 0 \end{pmatrix} \\ &= \sqrt{2} \sin \theta |0\rangle_{Anc} \otimes \begin{pmatrix} 1/\sqrt{2} \\ \pm 1/\sqrt{2} \end{pmatrix} + \sqrt{\cos 2\theta} |1\rangle_{Anc} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (1) \end{aligned}$$

<sup>2</sup>States written using the Fock notation  $|\cdot\rangle^F$  are called Fock states, see Section 4.

<sup>3</sup>The six-state scheme uses the three conjugate bases of the qubit space; namely, also  $|0_y\rangle = (|0_z\rangle + i|1_z\rangle)/\sqrt{2}$ , etc.

<sup>4</sup>By the term “forget” we mean that Bob's detection is unable to distinguish between several measured states.

<sup>5</sup>This entire process can be described in a compact way by using a POVM [39].

where  $|1\rangle_{Anc} \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{Anc}$ . This operation leads to a conclusive result with probability  $2\sin^2\theta$  (when the measured ancilla is  $|0\rangle_{Anc}$ ), and inconclusive result otherwise. It is simple to see that the same measurement can be done, *without the use of an ancilla*, if the states  $\theta_0$  and  $\theta_1$  are embedded at Alice's lab in a larger space  $M$ , e.g.  $M = H_3$ , using Bob's transformation

$$\begin{pmatrix} \cos\theta \\ \pm\sin\theta \\ 0 \end{pmatrix} \xrightarrow{\mathcal{U}} \begin{pmatrix} \sin\theta \\ \pm\sin\theta \\ \sqrt{\cos 2\theta} \end{pmatrix}. \quad (2)$$

In the general case, the space  $M$  might be very large, even infinite. Bob might use only parts of it, for his measurements.

A complication in performing security analysis is due to Bob's option to *both* use an ancilla and extend the space used by Alice. Our analysis in the following sections starts with the space extension only (Sections 2.3–2.4), and later on deals with the general case (Sections 2.5–2.6).

### 2.3 Bob's space, without an ancilla

Let us formulate the spaces involved in the protocol, as described above. Assume Alice uses the space  $H^A$  according to Definition 1, which is embedded in a (potentially larger) space  $M$ . Ideally, in the BB84 protocol, Bob would like to measure just the states in  $H^A$ , but in practice he usually can not do so. Each one of Alice's states  $|\psi_i\rangle_A$  is transformed by Bob's equipment into some pure<sup>6</sup> state  $|\psi_i\rangle_M \in M$ . The space which is spanned by those states contains all the information about Alice's states  $\{|\psi_i\rangle_A\}$ .

More important, Bob might be measuring un-needed subspaces of  $M$  which Alice's states do not span. For instance, examine the case where Bob uses detectors to measure the Fock states  $|1, 0\rangle^F$  and  $|0, 1\rangle^F$ . Bob is usually able to distinguish a loss (the state  $|0, 0\rangle^F$ ) or an error (e.g.  $|1, 1\rangle^F$ , one horizontal photon and one vertical photon), from the two desired states, but he cannot distinguish between other states containing multiple photons. This means that Bob measures a much larger subspace of the entire space  $M$ , but (inevitably) interprets outcomes outside  $H^A$  as legitimate states; e.g. the states  $|2, 0\rangle^F$ ,  $|3, 0\rangle^F$ , etc. are (mistakenly) interpreted as  $|1, 0\rangle^F$ . See further discussion in Section 4.3.

We denote Bob's setup (beam splitters, phase shifters, etc.) by the unitary operation  $\mathcal{U}_B$ , followed by a measurement; all these operations are operating on the space  $M$  (or parts of it). Bob might have several different setups (e.g. a different setup for the  $z$ -basis and for the  $x$ -basis). Let  $\mathbf{U}$  be the set of unitary transformations in all Bob's setups.

**Definition 3. [This definition is Temporary.]** Given a specific setup-transformation  $\mathcal{U}_j \in \mathbf{U}$ , let  $H^{B_j} \subseteq M$  be the subsystem actually measured by Bob, having  $K$  basis states  $\{|\phi_k\rangle_{B_j}\}_{k=0\dots K-1}$ . The set of **Bob's Measured Spaces** is the set  $\{H^{B_j}\}_{j=0\dots J-1}$  of  $J = |\mathbf{U}|$  spaces.

We have already seen that Bob might be measuring un-needed dimensions. On the other hand he might not measure certain subspaces of  $M$ , even when Alice's state might reach there. In either case, the deviation is commonly due to limitations of Bob's equipment.

### 2.4 The quantum space of the protocol, without an ancilla

The “quantum space of the protocol” (QSoP) is in fact Alice's *extended* space, taking into consideration its *extensions* due to Bob's measurements. The security analysis of a protocol depends on the space  $H^{B^{-1}}$  defined below.

---

<sup>6</sup>The case in which Bob transposes the state into a mixed state is a special case of the analysis done in Section 2.5. For the notion of mixed states or quantum mixture see [37, 39].

**Definition 4. [This definition is Temporary.]** The reversed space  $H^{B^{-1}}$  is the Hilbert space spanned by the states  $\mathcal{U}_j^{-1}(|\phi_k\rangle_{B_j})$ , for each possible setup  $\mathcal{U}_j \in \mathbf{U}$ , and for each basis state  $|\phi_k\rangle_{B_j}$  of the appropriate  $H^{B_j} \subseteq M$ .

The Space  $H^{B^{-1}}$  usually resides in a larger space than  $H^A$ . For instance, using photons, the ideal space  $H^A$  consists of two modes with 2 basis states, see Section 4. Now  $H^{B^{-1}}$  could have an infinite space in each mode, but also could have more modes.

In order to derive the quantum space of the protocol we need to define the way Alice's space is extended according to  $H^{B^{-1}}$ , for this simple case where Bob does not add an ancilla. In this case, the space  $H^{B^{-1}}$  simply extends Alice's space to yield the QSoP via  $H^P = H^A + H^{B^{-1}}$ . Formally speaking

**Definition 5. [This definition is Temporary.]** The Quantum Space of the Protocol,  $H^P$ , is the space spanned by the basis states of the space  $H^A$  and the basis states of the space  $H^{B^{-1}}$ .

If Alice's realistic space is fully measured by Bob's detection process, then  $H^A$  is a subspace of  $H^{B^{-1}}$ , hence  $H^P = H^{B^{-1}}$ .

## 2.5 Bob's space (general case)

In the general case, one must consider Bob's option to add an ancilla during his measurement process. This addition causes a considerable difficulty in analyzing a protocol, however it is often an inherent part of the protocol, and can not be avoided. We denote the added ancilla as the state  $|0\rangle_{B'}$  that resides in the space  $H^{B'}$ .

**Definition 6.**  $M'$  is the space that includes the physical space used by Alice as defined in Definition 2, in addition to Bob's ancilla,  $M' = M \otimes H^{B'}$ .

Bob measures a subspace of the space  $M'$ , so the (permanent) definitions of his measured spaces  $H^{B_j}$  and the reversed space  $H^{B^{-1}}$  should be modified accordingly.

**Definition 7.** Given a specific setup-transformation  $\mathcal{U}_j \in \mathbf{U}$  let  $H^{B_j} \subseteq M'$  be the subsystem actually measured by Bob, having  $K$  basis states  $\{|\phi_k\rangle_{B_j}\}_{k=0\dots K-1}$ . The set of **Bob's Measured Spaces**, is the set  $\{H^{B_j}\}_{j=0\dots J-1}$  of  $J = |\mathbf{U}|$  spaces.

## 2.6 The quantum space of the protocol (general case)

The quantum space of the protocol is still Alice's *extended* space, while considering its *extensions* due to Bob's measurements. Yet, the added ancilla makes things much more complex. The security analysis of a protocol depends now *not* on the space  $H^{B^{-1}}$  defined below, but on a (potentially *much larger*) space obtained from it by tracing-out Bob's ancilla. As before, we first define the reversed space.

**Definition 8.** The reversed space  $H^{B^{-1}}$  is the Hilbert space spanned by the states  $\mathcal{U}_j^{-1}(|\phi_k\rangle_{B_j})$ , for each possible setup  $\mathcal{U}_j \in \mathbf{U}$ , and for each basis state  $|\phi_k\rangle_{B_j}$  of the appropriate  $H^{B_j} \subseteq M'$ .

Once a basis state of one of Bob's measured spaces  $|\phi_k\rangle_{B_j}$  is reversed by  $\mathcal{U}_j^{-1}$  we result with a state that might, partially, reside in Bob's ancillary space  $H^{B'}$ . Since Eve has no access to this space<sup>7</sup> it must be traced-out (separated out), for deriving the QSoP. Let us redefine the QSoP given the addition of the ancilla:

<sup>7</sup>Giving this space to Eve (for getting an upper bound on her information), might be easier to analyze, but is usually not possible since it would give her too much power, making the protocol insecure.

**Definition 9. The Quantum Space of the Protocol,  $H^P$ ,** is the space spanned by (a) the basis states of the space  $H^A$ ; and (b) the states  $\text{Tr}_{\text{Bob}}[\mathcal{U}_j^{-1}(|\phi_k\rangle_{B_j})]$ , (namely, after tracing out Bob), for each possible setup  $\mathcal{U}_j \in \mathcal{U}$ , and for each basis state  $|\phi_k\rangle_{B_j}$  of the appropriate space  $H^{B_j}$ .

Whenever  $\mathcal{U}_B$  entangles Bob's ancilla with the system sent from Alice, tracing out Bob's ancilla after performing  $\mathcal{U}_B^{-1}$  might cause an increase of the QSoP to the dimension of Bob's ancillary space. For instance, assume Alice's state is embedded in an  $n$ -qubit space to which Bob adds an ancilla of  $n$ -qubits and performs a unitary transformation  $\mathcal{U}$ , such that for one state measured by Bob,  $|\Psi\rangle_B \xrightarrow{\mathcal{U}^{-1}} \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} |k\rangle_P |k\rangle_{B'}$ . Tracing out Bob from this state yields the maximally mixed state  $\rho_P = \frac{1}{2^n} \sum_{k=0}^{2^n-1} |k\rangle\langle k|$ , so that in this example the whole  $n$ -qubits space is spanned.

### 3 The Quantum Space Attack

#### 3.1 Eavesdropping on qubits

When Alice and Bob use qubits, in theoretical QKD, Eve can attack the protocol in many ways. In her simplest attack, the so-called “measure-resend attack”, Eve performs any measurement (of her choice) on the qubit, and accordingly decides what to send to Bob.

A generalization of that attack is the “translucent attack”, in which Eve attaches an ancilla, in an initial state  $|0\rangle_E$  (and in any dimension she likes), and entangles the ancilla and Alice's qubit, using  $|0\rangle_E |i\rangle_A \rightarrow \sum_{j=0}^1 |E_{ij}\rangle_E |j\rangle_A$  where  $|i\rangle_A$  is a basis for Alice's qubit, and Eve's states after the unitary transformation are  $|E_{ij}\rangle_E$ . Using this transformation one can define the most general “individual-particle attack” [19, 20], and also the most general “collective attack” [9, 8]. In the individual-particle attack Eve delays the measurement of her ancilla till after learning anything she can about the qubit (e.g., its basis), while in the collective attack Eve delays her measurements further till she learns anything she can about *all* the qubits (e.g., how the final key is generated from the obtained string of shared bits), so she attacks directly the *final key*.

The most general attack that Eve could perform on the channel is to attack all those qubits transmitted from Alice to Bob, using *one* large ancilla. This is the “joint attack”. Security, in case Eve tries to learn a maximal information on the final key, was proven in [46, 35, 6, 42, 7] via various methods. The attack's unitary transformation is written as before, but with  $i$  a binary string of  $n$  bits, and so is  $j$ ,  $|0\rangle_E |i\rangle_A \rightarrow \sum_{j=0}^{2^n-1} |E_{ij}\rangle_E |j\rangle$ .

#### 3.2 Eavesdropping on the quantum space of the protocol

By replacing the qubit space  $H_2$  by Alice's realistic “qubit” in the space  $H^A$ , and by defining Eve's attack on the entire space of the protocol  $H^P$ , we can generalize each of the known attacks on theoretical QKD to a “quantum space attack” (QSA). We can easily define now Eve's most general *individual-transmission QSA* on a realistic “qubit”, which generalizes the individual-particle attack earlier described. Eve prepares an ancilla in a state  $|0\rangle_E$ , and attaches it to Alice's state, but actually her ancilla is now attached to the entire QSoP. Eve performs a unitary transformation  $\mathcal{U}_E$  on the joint state. If Eve's attack is only on  $H^A$ , we write the resulting transformation on any basis state of  $H^A$ ,  $|i\rangle_A$ , as  $|0\rangle_E |i\rangle_A \rightarrow \sum_j |E_{ij}\rangle_E |j\rangle_A$ , where the sum is over the dimension of  $H^A$ . The Photon-Number-Splitting attack (see Section 5.1) is an example for such an attack. The most general individual-transmission QSA is based on a translucent QSA on the QSoP,

$$|0\rangle_E |i\rangle_P \rightarrow \sum_j |E_{ij}\rangle_E |j\rangle_P, \quad (3)$$

where the sum is over the dimension of  $H^P$ . The subsystem in  $H^P$  is then sent to Bob while the rest (the subsystem  $H^E$ ) is kept by Eve. We write the transformation on any basis state of  $H^P$ ,  $|i\rangle_P$ , but note that

it is sufficient to define the transformation on the different states in  $H^A$ , namely for all states of the form  $|i\rangle_A$ , since other states of the QSoP are never sent by Alice (any other additional subsystem of the QSoP is necessarily at a known state when it enters Eve's transformation).

Attacks that are more general than the *individual transmission QSA*, the *collective QSA* and the *joint QSA*, can now be defined accordingly. In the most general collective QSA, Eve performs the above translucent QSA on many (say,  $n$ ) realistic “qubits” (potentially a different attack on each one, if she likes), waits till she gets all data regarding the generation of the final key, and she then measures all the ancillas together, to obtain the optimal information on the final key or the final secret. The most general attack that Eve could perform on the channel is to attack all those realistic “qubits” transmitted from Alice to Bob, using *one* large ancilla. This is the “joint QSA”. The attack's unitary transformation is written as before, but with  $i$  a string of  $n$  digits rather than a single digit (digits of the relevant dimension of  $H^P$ ), and so is  $j$ ,

$$|0\rangle_E |i\rangle_{P^{\otimes n}} \rightarrow \sum_{j=0}^{|H^P|^n - 1} |E_{ij}\rangle_E |j\rangle_{P^{\otimes n}}. \quad (4)$$

Eve measures the ancilla, after learning all classical information, to obtain the optimal information on the final key or the final secret. As before, it is sufficient to define the transformation on the different input states from  $(H^A)^{\otimes n}$ .

We would like to emphasize several issues: 1.– When analyzing specific attacks, or when trying to obtain a limited security result, it is always legitimate to restrict the analysis to the relevant (smaller) subspace of the QSoP, for simplicity, e.g., to  $H^A$ , or to  $H^{B^{-1}}$ , etc. 2.– Any bi-directional protocol will have a much more complicated QSoP, thus it might be extremely difficult to analyze any type of QSA (even the simplest ones) on such protocols. This remark is especially important since bi-directional protocols play a very important role in QKD, since they appear in many interesting protocols such as the plug-and-play [33], the ping-pong [10], and the classical Bob [11] protocols. Specifically they provided (via the plug-and-play) the only commercial QKD so far [48, 49]. 3.– It is well known that the collective or joint attack is only finished after Eve gets all quantum and classical information, since she delays her measurements till then [9, 8, 6, 35, 7]; if she expects more information, she better wait and attack the final secret rather than the final key; it is important to notice that if the key will be used to encode quantum information (say, qubits) then the quantum-space of the protocol will require a modification, potentially a major one; It is interesting to study if this new notion of QSoP has an influence on analysis of such usage of the key as done (for the ideal qubits) in [4].

## 4 Photonic Quantum Space Attacks

### 4.1 Photons as quantum-information carriers

Since most of the practical QKD experiments and products are done using photons, in this section we demonstrate our QSoP and QSA definitions and methods via photons. Our analysis uses the Fock-Space<sup>8</sup> notations for describing photonic quantum spaces. For clarity, states written using the Fock notation are denoted with the superscript ‘ $F$ ’, e.g.  $|0\rangle^F$ ,  $|3\rangle^F$ , and  $|0, 3, 1\rangle^F$ .

A photon can not be treated as a quantum system in a straightforward way. For instance, unlike dust particles or grains of sand, photons are indistinguishable particles, meaning that when a couple of photons are interacting, one cannot define the evolution of the specific particle, but rather describe the whole system.

Let us examine a cavity, for instance. It can contain photons of specific wavelenghtes ( $\lambda_1$ ,  $\lambda_2$ , etc.) and the energy of a photon of wavelength  $\lambda$  is directly proportional to  $1/\lambda$ . While one cannot distinguish between

<sup>8</sup>A description of the Fock space and Fock notations can be found in various quantum optic books, e.g. [41].

photons of the same wavelength, one can distinguish between photons of different wavelengths. Therefore, it is convenient to define distinguishable “photonic modes”, such that each wavelength corresponds to a specific mode (so a mode inside a cavity can be denoted by its wavelength), and then count the number of photons in each mode. If a single photon in a specific mode carries some unit of energy, then  $n$  such photons of the same wavelength carry  $n$  times that energy. If the cavity is at its ground (minimal) energy level, we say that there are “no photons” in the cavity and denote the state as  $|0\rangle^F$ —the vacuum state. The convention is to denote only those modes that are potentially populated, so if we can find  $n$  photons in one mode, and no photons in any other mode, we write,  $|n\rangle^F$ . If two modes are populated by  $n_a$  and  $n_b$  photons, and all other modes are surely empty, we write  $|n_a, n_b\rangle^F$  (or  $|m, n\rangle_{ab}^F$ ). When there is no danger of confusion, and the number of photons per mode is small (smaller than ten), we just write  $|mn\rangle^F$  for  $m$  photons in one mode and  $n$  in the other. In addition to its wavelength, a photon also has a property called polarization, and a basis for that property is, for instance, the horizontal and vertical polarizations mentioned earlier. Thus, two modes (in a cavity) can also have the same energy, but different polarizations.

Outside a cavity photons travel with the speed of light, say from Alice to Bob, yet modes can still be described, e.g., by using “pulses” of light [14]. The modes can then be distinguished by different directions of the light beams (or by different paths), or by the timing of pulses (these modes are denoted by non-overlapping time-bins), or by orthogonal polarizations.

A proper description of a photonic qubit is commonly based on using two modes ‘ $a$ ’ and ‘ $b$ ’ which are populated by exactly a single photon, namely, a photon in mode  $a$ , so the state is  $|10\rangle_{ab}^F$ , or a photon in mode  $b$ , so the state is  $|01\rangle_{ab}^F$ . However, a quantum space that consists of a single given photonic mode ‘ $a$ ’ is not restricted to a single photon, and can be populated by any number of photons. A basis for this space is  $\{|n\rangle_a^F\}$  with  $n \geq 0$ , so that the quantum space is infinitely large,  $H_\infty$ . Theoretically, a general state in this space is can be written as the superposition  $\sum_{n=0}^{\infty} c_n |n\rangle_a^F$ , with  $\sum_n |c_n|^2 = 1$ ,  $c_n \in \mathbb{C}$ . Similarly, a quantum space that consists of two photonic modes has the basis states  $|n_a, n_b\rangle^F$ , for  $n_a, n_b \geq 0$  and a general state is of the form  $\sum_{n_a, n_b=0}^{\infty} c_{n_a, n_b} |n_a, n_b\rangle^F$  with  $\sum_{n_a, n_b=0}^{\infty} |c_{n_a, n_b}|^2 = 1$ ,  $c_{n_a, n_b} \in \mathbb{C}$ . This quantum space is described as a tensor product of two “systems”  $H_\infty \otimes H_\infty$ .

Using *exactly* two photons in two different (and orthogonal) modes assists in clarifying the difference between photons and dust particles (or grains of sand): Due to the indistinguishability of photons, only 3 different states can exist (instead of 4):  $|20\rangle_{ab}^F$ ,  $|02\rangle_{ab}^F$  and  $|11\rangle_{ab}^F$ . The last state has one photon in mode ‘ $a$ ’ and another photon in ‘ $b$ ’, however, exchanging the photons is meaningless since one can never tell one photon from another.

A realistic model of a photon source (in a specific mode) is of a coherent pulse (a Poissonian distribution)

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

including terms that describe the possibility of emitting any number  $n$  of photons. As the number of photons increases beyond some number, the probability decreases, so it is common to neglect the higher orders. In QKD, experimentalists commonly use a “weak” coherent state (such that  $|\alpha| \ll 1$ ) and then terms with  $n \geq 3$  can usually be neglected. There is also a lot of research about sources that emit (to a good approximation) single photons, and then, again, terms with  $n \geq 3$  can usually be neglected.

## 4.2 Alice’s realistic photonic space

While the theoretical qubit lives in  $H_2$ , a realistic view defines the space actually used by Alice to be much larger. The possibility to emit empty pulses increases Alice’s realistic space into  $H_3$ , due to the vacuum state  $|00\rangle_{ab}^F$ . When Alice sends a qubit using two modes, using a weak coherent state (or a “single-photon” source), her realistic space,  $H^A$ , is embedded in  $H_\infty \otimes H_\infty$ . Terms containing more than two photons can be



neglected, so these are excluded from Alice’s space  $H^A$ . The appropriate realistic quantum space of Alice,  $H^A$ , is now a quhexit: the six-dimensional space spanned by  $\chi^6 = \{|00\rangle^F, |10\rangle^F, |00\rangle^F, |11\rangle^F, |20\rangle^F, |02\rangle^F\}$ . The PNS attack demonstrated in Section 5.1, is based on attacking this 6 dimensional space  $H^A$ . Note also that terms with more than two photons still appear in  $M$ , and thus could potentially appear in the QSoP (and then used by Eve).

At times, Alice’s realistic space is even larger, due to extra modes that are sent through the channel, and are not meant to be a part of the protocol. These extra modes might severely compromise the security of the protocol, since they might carry some vital information about the protocol. A specific QSA based on that flaw is the “tagging attack” (Section 5.2). Note that even if Alice uses exactly two modes, the quantum space  $M$  where  $H^A$  is embedded, certainly contains other modes as well.

### 4.3 Extensions of the photonic space; the QSoP

Let us discuss Bob’s measurement of photonic spaces. There are (mainly) two types of detectors that can be used. The common detector can not distinguish a single photon from more than one photon (these kind of detectors are known as *threshold detectors*). The Hilbert space where Bob’s measurement is defined is infinite<sup>9</sup>, since a click in the detector tells Bob that the number of photons occupying the mode is “not zero” i.e. the detector clicks when  $|n\rangle^F$  is detected, for  $n \geq 1$ . This means that Bob measures the state  $|0\rangle^F$ , or he measures  $|1\rangle^F, |2\rangle^F, \dots$  but then “forgets” how many photons were detected. Bob might severely compromise the security, since he inevitably interprets a measurement of a state containing multiple photons as the “legal” state that contains only a single photon. An attack based on a similar limitation is the “Trojan-Pony” attack described below, in Section 5.3. In order to avoid false interpretations of the photon number reaching the detector, Bob could use an enhanced type of detector known as the *photon-number resolving detector* or a *counter* (which is still under development). This device distinguishes a single photon from  $n \geq 2$  photons, hence any eavesdropping attempt that generates multi-photon states can potentially be noticed by Bob. A much enhanced security can be achieved now, although the QSoP is infinite also in this case, due to identifying correctly the legitimate state  $|1\rangle^F$ , from various legitimate states.

The number of modes in the QSoP depends on Bob’s detectors as well. Bob commonly increases the number of measured modes by “opening” his detector for more time-bin modes or more frequency modes. For instance, suppose Bob is using a detector whose detection time-window is quite larger than the width of the pulse used in the protocol, since he does not know when exactly Alice’s pulse might arrive. The result is an extension of the space used by Alice, so that the QSoP includes the subspace of  $M$  that contains all these measured modes. When a single detector is used to measure more than one mode *without distinguishing them*, the impact on the security might be severe, see the “Fake state” attack (Section 5.4).

In addition to the known attacks described in the following subsection, a new QSA is analyzed in Section 6, where we examine the more general case of QSA, in which Bob adds an ancilla during the process.

## 5 Known Attacks as Quantum-Space Attacks

All known attacks can be considered as special cases of the Quantum-Space Attack. In this section we show a description of several such attacks using QSA terms. For each and every attack we briefly describe the specific protocol used, the quantum space of the protocol, and a realization of the attack as a QSA.

---

<sup>9</sup> In practice, that space is as large as Eve might wish it to be. We can ignore the case where Eve uses too many photons so that the detector could burn due to the high energy, since it is not in Eve’s interest. Thus, in some of the analyses below we replace  $\infty$  by some large number  $L$ .

## 5.1 The photon number splitting attack [13]

**The Protocol.** Consider a BB84 protocol, where Alice uses a “weak pulse” laser to send photons in two modes corresponding to the vertical and horizontal polarizations when using the  $z$  basis (the diagonal polarizations then relate to using the  $x$  basis). Bob uses a device called a Pockel cell to rotate the polarization (by  $45^\circ$ ) for measuring the  $x$  basis, or performs no rotation if measuring the  $z$  basis. The measurement of the state is then done using two detectors and a “polarization beam splitter” that passes the first mode to one detector and the second mode to the other detector (for a survey of polarization-based QKD experiments, see [23, 17]).

**The Quantum Space of the Protocol.** Every pulse sent by Alice is in one of four states, each in a superposition of the 6 orthogonal states  $\chi^6 = \{|00\rangle^F, |10\rangle^F, |01\rangle^F, |11\rangle^F, |20\rangle^F, |02\rangle^F\}$ , where the space used by Alice is  $H^A = H_6$ . Bob uses two setups,  $\mathcal{U}_{B_z} = I$  for the  $z$  basis, and  $\mathcal{U}_{B_x}$  for the  $x$  basis, which is more complex and described in Appendix A.1.

The detectors used by Bob cannot distinguish between modes having single photon and multiple photons. Each one of his two detectors measures the basis elements  $\{|n\rangle^F\}$  for  $n \geq 0$  (of the specific mode directed to that specific detector), where Bob interprets the states  $\{|n\rangle^F\}$  with  $n > 1$  as measuring the state  $|1\rangle^F$  of the same mode. Bob’s measured space  $H^B$  is thus infinite and spanned by the states  $\{|mn\rangle^F\}$  for  $m, n \geq 0$ . The QSoP  $H^P$  is equal to  $H^{B_z} (= H^{B_x})$  since performing  $\mathcal{U}^{-1}$  does not change the dimensionality of the spanned space (in both setups).

**The Attack.** Eve measures the number of photons in the pulse, using non-demolition measurement. If she finds that the number of photons is  $\geq 1$ , she blocks the pulse and generates a loss. In the case she finds that the pulse consists of 2 photons, she splits one photon out of the pulse and sends it to Bob, keeping the other photon until the bases are revealed, thus getting full information of the key-bit. Eve sends the eavesdropped qubits to Bob via a lossless channel so that Bob will not notice the enhanced loss-rate. As is common in experimental QKD, Bob is willing to accept a high loss-rate (he does not count losses as errors), since most of Alice’s pulses are empty. See the precise mathematical description of this attack in Appendix A.

## 5.2 The tagging attack (based on [24])

**The Protocol.** Consider a BB84 QKD protocol in which Alice sends an enlarged state rather than a qubit. This state contains, besides the information qubit, a *tag* giving Eve some information about the bit. The tag can, for example, tell Eve the basis being used by Alice. For a potentially realistic example, let the tag be an additional qutrit indicating if Alice used the  $x$ -basis, or the  $z$ -basis, or whether the basis is *unknown*: whenever Alice switches basis, a single photon comes out of her lab prior to the qubit-carrying pulse, telling the basis, say using the states  $|10\rangle_{tag}^F$  and  $|01\rangle_{tag}^F$ , and when there is no change of basis, what comes out prior to the qubit is just the vacuum  $|00\rangle_{tag}^F$ .

**The Quantum Space of the Protocol.** In this example, Alice is using the space  $H^A = H_2 \otimes H_{tag} = H_2 \otimes H_3$ . Bob, unaware of the enlarged space used by Alice, expects and receives only the subspace  $H_2$ . We assume that Bob ideally measures this space with a single setup  $\mathcal{U}_B = I$ , therefore  $H^B = H_2$ . Since Bob’s setup does not change the space,  $H^{B^{-1}} = H_2$  as well. However, the tag is of a much use to Eve, and indeed the QSoP following Definition 5, defined to be  $H^P = H_2 \otimes H_{tag}$ .

**The Attack.** Eve uses the tag in order to retrieve information about the qubit without inducing error (e.g. via cloning the qubit in the proper basis). The attack is then an intercept-resend QSA. We mention that this attack is very similar to a side-channel cryptanalysis of classic cryptosystems.

**A Short Summery.** It can be seen that the PNS attack described above is actually a special case of the tagging attack, where the *tag* in that case is in fact another copy of the transmitted qubit. This copy

is kept by Eve until the bases are revealed, then it can be measured so the the key-bit value is exposed with certainty. Both those QSA attacks are based on the fact that Alice (realistic) space is larger than the theoretical one. Although in the PNS example, the QSoP is further extended due to Bob's measurement, the attack is not based on that extension but on the fact that  $H^A$  is larger than  $H_2$ . In the following attacks Bob's measurements cause the enlargement of the QSoP, allowing Eve to exploit the larger QSoP for her attack.

### 5.3 The Trojan-pony attack [24, 26]

In Trojan-pony attacks Eve modifies the state sent to Bob in a way that gives her information. In contrast to a "Trojan-horse" that goes in-and-out of Bob's lab, the "pony" only goes in, therefore, it is not considered an attack on the lab, but only on the channel. We present here an interesting example [24].

**The Protocol.** Assume a polarization-encoded BB84 protocol, in which Alice is ideal, namely, sending perfect qubits ( $H^A = H_2$ ). However, Bob uses realistic threshold detectors that suffer from losses and dark counts, and that cannot distinguish between one photon and  $k$  photons for  $1 < k < L$ . In order to be able to "prove" security, for a longer distance of transmission Bob wants to keep the error-rate low although the increase of dark counts' impact with the distance [13]. Therefore, Bob assumes that Eve has no control over dark counts, and whenever both detectors click, Alice and Bob agree to consider it as a *loss* since it is outside of Eve's control (i.e. the QSoP is falsely considered to be  $H_2$ ). Namely, they assume that *an error* occurs only when Bob measures in the right basis, and only one detector clicks, (which is the detector corresponding to the wrong bit-value).

**The Quantum Space of the Protocol.** Same as in Section 5.1, Bob's measured spaces  $H^{B_z}$ ,  $H^{B_x}$ , the reversed space  $H^{B^{-1}}$  as well as the QSoP  $H^P$ , are merely the spaces describing two modes (with up to  $L$  photons),  $H_L \otimes H_L$ . Bob's detectors cannot distinguish between receiving a single-photon pulse from a multi-photon pulse, so his measurement is properly described as a projection of the received state onto the space containing  $\{|ij\rangle^F\}$  followed by "forgetting" the exact result, and keeping only one of three results: " $\{10\} \equiv$  detector-1 clicks", " $\{01\} \equiv$  detector-2 clicks", and else it is  $\{00\}$ , a "loss". In formal, *generalized-measurements* language (called POVM, see [39, 37]) these three possible results are written as:  $\{10\} \equiv \sum_{k=1}^{L-1} |k0\rangle^F \langle k0|$ ,  $\{01\} \equiv \sum_{k=1}^{L-1} |0k\rangle^F \langle 0k|$ ,  $\{00\} \equiv |00\rangle^F \langle 00| + \sum_{k_1, k_2=1}^{L-1} |k_1 k_2\rangle^F \langle k_1 k_2|$ , and their sum is the identity matrix.

**The Attack.** Eve's attack is the following: (a) Randomly choose a basis (b) Measure the arriving qubit in that specific chosen basis (c) Send Bob  $m$ -photons identical to the measured qubit, where  $m \gg 1$ . Obviously, when Eve chooses the same basis as Alice and Bob then Bob measures the exact value sent by Alice, and Eve gets full information. Otherwise, both of his detectors click, implying a "loss", except for a negligible probability,  $\approx 2^{(-m+1)}$ , thus Eve induces no errors. The main observation of this measure-resend QSA is that treating a count of more than a single photon as a loss, rather than as an error, is usually not justified. A second conclusion is that letting Bob use counters instead of threshold detectors (to distinguish a single photon from multiple photons), together with treating any count of more than one photon as an error, could be vital for proving security against QSA. The price is that dark counts put severe restrictions on the distance to which communication can still be considered secure, as suggested already by [13].

### 5.4 The fake-state attack (based on [34, 31])

**The Protocol.** In this example, we examine a polarization encoded BB84 protocol, and an ideal Alice ( $H^A = H_2$ ). This time Bob's detectors are imperfect so that their detection windows do not fully overlap, meaning that there exist times in which one detector is blocked (or it has a low efficiency), while the other detector is still regularly active. Thus, if Eve can control the precise timing of the pulse, she can control whether the photon will be detected or lost. The setup is built four detectors and a rotating mirror (since Bob does not want to spend money on a Pockel cell (polarization rotator), he actually uses 2 fixed different

setups). Using the rotating mirror Bob sends the photon into a detection setup for basis  $z$  or a detection setup for basis  $x$ . Suppose the two detection setups use slightly different detectors, or slightly different delay lines, or slightly different shutters, and Eve is aware of this (or had learnt it during her past attacks on the system). For simplicity, we model the non-overlapping detection windows, as additional two modes, one slightly prior to Alice’s intended mode (the pulse), and one right after it.

**The Quantum Space of the Protocol.** The original qubit is sent in a specific time-bin  $t_0$  (namely,  $H^A = H_2$ ). The setup  $\mathcal{U}_Z$  is a set of two detectors and a polarized beam splitter, separating the horizontal and the vertical modes to the detectors, where  $\mathcal{U}_x$  separate the diagonal modes into a set of two (different) detectors. Let the detectors for one basis, say  $z$ , be able to measure a pulse arriving at  $t_0$  or  $t_1$ , while the detectors for the other basis ( $x$ ) measure pulses arriving at  $t_{-1}$  or  $t_0$ .

For simplicity, we degenerate the space to contain one or less photons<sup>10</sup>, so that  $H^{B_z}$  is  $H_5$ , i.e. two possible time-bins consisting each of two (polarization) modes of one or less photons. The measured space of the  $x$ -setup has two possible time-bins and two possible polarization modes, thus  $H^{B_x} = H_5$  as well, however, the two time-bins for this setup are  $t_0$  and  $t_1$ . Following Definition 4 we get that the reversed space  $H^{B^{-1}}$  contains three time-bins ( $t_{-1}$ ,  $t_0$  and  $t_1$ ) with two polarization modes in each, therefore  $H^{B^{-1}} = H_7$ , under the single-photon assumption. The QSoP, following Definition 5 equals  $H^{B^{-1}}$  since  $H^A \subset H^{B^{-1}}$ .

**The Attack.** Eve exploits the larger space by sending “fake” states using the external time bins ( $t_{-1}$  and  $t_1$ ). Eve randomly chooses a basis, measures the qubit sent by Alice, and sends Bob the same polarization state she found, but at  $t_{-1}$  if she has used the  $x$  basis, or at  $t_1$  if she has used the  $z$  basis. Since no ancilla is kept by Eve, this is an intercept-resend QSA.

Bob will get the same result as Eve if he uses the same basis, or a loss otherwise. The mathematical description of the attack is as follows: Eve can generate superpositions of states of the form  $|V_{t_{-1}}H_{t_{-1}}V_{t_0}H_{t_0}V_{t_1}H_{t_1}\rangle^F$ , where the index  $\{H, V\}$  denotes this mode has Vertical or Horizontal polarization, and its subscript denotes the time-bin of the mode. Eve’s measure-resend attack is described as measuring Alice’s qubit in the  $x$  basis, creating a new copy of the measured qubit, and performing the transformation  $(|001000\rangle^F \rightarrow |100000\rangle^F)$ ;  $(|000100\rangle^F \rightarrow |010000\rangle^F)$  or as performing a measurement in the  $z$  basis, and performing the transformation  $(|001000\rangle^F \rightarrow |000010\rangle^F)$ ;  $(|000100\rangle^F \rightarrow |000001\rangle^F)$  on the generated copy.

**A short summary** We see that Eve can “force” a desired value (or a loss) on Bob, thus gaining all the information while inducing no errors (but increasing the loss rate). Bob can use a shutter to block the irrelevant time-bins but such a shutter could generate a similar problem in the frequency domain. This attack is actually a special case of the Trojan-pony attack, in which the imperfections of Bob’s detectors allow Eve to send states that will be un-noticed unless the measured basis equals to Eve’s chosen basis.

## 6 Interferometric BB84 and 6-state Protocols

In order to demonstrate the power of QSA, and to see its advantages, this section presents a partial security analysis of some interferometric BB84 and 6-state schemes. Interferometric schemes are more common than any other type of implementation in QKD experiments [43, 32, 23, 18, 17, 33] and products [48, 49]. In this section we define the specific equipment used by Bob, and we formulate  $\mathcal{U}_B$  and Bob’s measurements. We then find the spaces  $H^A$ ,  $H^{B_j}$ ,  $H^{B^{-1}}$  and the QSoP,  $H^P$ . Finally, we demonstrate a novel attack which is found to be very successful against a specific variant of the BB84 interferometric scheme; this specific QSA, which we call the “reversed-space attack”, is designed using the tools developed in Sections 2 and 3.

<sup>10</sup>As mentioned above, this is used for non-security proof, and is not a legitimate assumption for proving unconditional security, where the three time-modes should be considered as  $H_L \otimes H_L \otimes H_L$ .

## 6.1 Bob's equipment

We begin with a description of interferometric (BB84 and six-state) schemes, which is based on sending phase-encoded qubits arriving in two time-separated modes [43, 32]. Alice encodes her qubit using two time-bins  $t'_0$  and  $t'_1$ , where a photon in the first mode,  $|10\rangle_{t'_0 t'_1}^F$ , represents the state  $|0_z\rangle$ , and a photon in the other mode,  $|01\rangle_{t'_0 t'_1}^F$ , represents  $|1_z\rangle$ . The BB84 protocol of [43, 32] (and many others) uses the  $x$  and  $y$  bases, meaning that Alice (ideally) sends one of the following four states:  $|0_x\rangle = (|10\rangle_{t'_0 t'_1}^F + |01\rangle_{t'_0 t'_1}^F)/\sqrt{2}$ ;  $|1_x\rangle = (|10\rangle_{t'_0 t'_1}^F - |01\rangle_{t'_0 t'_1}^F)/\sqrt{2}$ ;  $|0_y\rangle = (|10\rangle_{t'_0 t'_1}^F + i|01\rangle_{t'_0 t'_1}^F)/\sqrt{2}$ ; and  $|1_y\rangle = (|10\rangle_{t'_0 t'_1}^F - i|01\rangle_{t'_0 t'_1}^F)/\sqrt{2}$ .

Bob uses an interferometer built from two beam splitters with one short path and one long path (Figure 1). A pulse of light travels through the short arm of the interferometer in  $T_{\text{short}}$  seconds, and through the long arm in  $T_{\text{long}} = T_{\text{short}} + \Delta T$  seconds, where  $\Delta T$  is also *precisely* the time separation between the two arriving modes of the qubit,  $\Delta T = t'_1 - t'_0$ . A controlled phase shifter  $P_\phi$ , is placed in the long arm of the interferometer. It performs a phase shift by a given phase  $\phi$ , i.e.  $P_\phi(|\psi\rangle) = e^{i\phi}|\psi\rangle$ . The phase shifter is set to  $\phi = 0$  ( $\phi = \pi/2$ ) when Bob measures the  $x$  ( $y$ ) basis. Each beam splitter interferes two input arms (modes 1, 2) into two output arms (modes 3, 4), in the following way (for a single photon):  $|10\rangle_{1,2}^F \mapsto \frac{1}{\sqrt{2}}|10\rangle_{3,4}^F + \frac{i}{\sqrt{2}}|01\rangle_{3,4}^F$ , and  $|01\rangle_{1,2}^F \mapsto \frac{i}{\sqrt{2}}|10\rangle_{3,4}^F + \frac{1}{\sqrt{2}}|01\rangle_{3,4}^F$ . The photon is transmitted/reflected with a probability of 50%; The transmitted part keeps the same phase as the incoming photon, while the reflected part gets an extra phase of  $e^{i\pi/2}$ , if it carries a single photon. When a single mode, carrying at least a single photon, enters a beam splitter from one arm, and nothing enters the other input arm, we must consider the other entry to be an additional mode (an ancilla) in a vacuum state.

When a single mode (carrying one or more photons) enters the interferometer at time  $t'_0$ , see Figure 1, it yields two modes at time  $t_0$  due to traveling through the short arm, and two modes at time  $t_1$  due to traveling through the long arm. Those four output modes are: times  $t_0, t_1$  in the 's' (straight) arm of the interferometer, and times  $t_0, t_1$  in the 'd' (down) arm. A basis state in this Fock space is then  $|n_{s_0}, n_{s_1}, n_{d_0}, n_{d_1}\rangle^F$ . In the case of having that single mode carrying exactly a single photon, the transformation, which requires three additional empty ancillas<sup>11</sup>, is  $|1\rangle_{t'_0}^F |000\rangle^F \mapsto (|1000\rangle^F - |0100\rangle^F + i|0010\rangle^F + i|0001\rangle^F)/2$ . Note that a pulse which is sent at a different time (say,  $t'_x$ ) results in the same output state, but with the appropriate delays, i.e.

$$|1\rangle_{t'_x}^F |000\rangle^F \mapsto (|1000\rangle^F - |0100\rangle^F + i|0010\rangle^F + i|0001\rangle^F)/2, \quad (5)$$

where the resulting state is defined in the Fock space whose basis states are  $|n_{s_x}, n_{s_{x+1}}, n_{d_x}, n_{d_{x+1}}\rangle$ .

Let us now examine any superposition of two modes ( $t'_0$  and  $t'_1$ ) that enter the interferometer one after the other, with exactly the same time difference  $\Delta T$  as the difference lengths of the arms. The state evolves in the following way (see Appendix B.2):

$$\begin{aligned} \cos \theta |10\rangle_{t'_0 t'_1}^F |0000\rangle^F + \sin \theta e^{i\varphi} |01\rangle_{t'_0 t'_1}^F |0000\rangle^F \mapsto \\ \left( \cos \theta |100000\rangle_B^F + (-\cos \theta e^{i\phi} + \sin \theta e^{i\varphi}) |010000\rangle_B^F - \sin \theta e^{i(\varphi+\phi)} |001000\rangle_B^F \right. \\ \left. + i \cos \theta |000100\rangle_B^F + i(\cos \theta e^{i\phi} + \sin \theta e^{i\varphi}) |000010\rangle_B^F + i \sin \theta e^{i(\varphi+\phi)} |000001\rangle_B^F \right) / 2 \end{aligned} \quad (6)$$

describing the evolution for any possible BB84 state sent by Alice ( $|0_x\rangle, |1_x\rangle, |0_y\rangle, |1_y\rangle$ ) determined by the value of  $\varphi = 0, \pi, \frac{\pi}{2}, \frac{3\pi}{2}$  respectively, when  $\theta = \frac{\pi}{4}$ . As a result of this precise timing, these two modes are transformed into a superposition of 6 possible modes (and not 8 modes) at the outputs, due to interference at the second beam splitter. Only four vacuum-states ancillas (and not six) are required for that process. The resulting 6 modes are  $t_0, t_1, t_2$  in the 's' arm and in the 'd' arm of the interferometer. Denote this Fock space as  $H^B$ , with basis elements  $|n_{s_0}, n_{s_1}, n_{s_2}, n_{d_0}, n_{d_1}, n_{d_2}\rangle_B^F$ .

<sup>11</sup> See a brief description in Appendix B.1.

The measurement is performed as follows: Bob opens his detectors at time  $t_1$  in both output arms of the interferometer. A click in the “down” direction means measuring the bit-value 0, while a click in the “straight” direction means 1. The other modes are commonly considered as a loss (they are not measured) since they give an inconclusive result regarding the original qubit. We refer this BB84 variant as “ $xy$ -BB84”.

One might want to use the  $z$  basis in his QKD protocol (using  $\varphi = 0$ , and  $\theta = 0$  or  $\theta = \frac{\pi}{2}$ ), for instance, in order to avoid the need for a controlled phase shifter or for another equipment-related reason, or in order to perform “QKD with classical Bob” [11]. A potentially more important reason might be to perform the 6-state QKD [15, 3, 29] protocol, due to its improved immunity against errors (27.4% errors versus only 20% in BB84 [16]). A possible and easy to implement variant for realizing a measurement in the  $z$  basis is the following: Bob uses the setup  $\mathcal{U}_{B_x}$  (i.e. he sets  $P_\phi$  to  $\phi = 0$ ), and opens his detectors at times  $t_0$  and  $t_2$ , corresponding to the bit-values 0 and 1 respectively (See Equation (6)). Unfortunately, technological limitations, e.g. of telecommunication wavelength (IR) detectors, might make it difficult for Bob to open his detectors for more than a single detection window per pulse. Bob could perform a measurement of *just* the states  $\{|000100\rangle_B^F, |001000\rangle_B^F\}$ , opening the  $d$  arm detector at time  $t_0$  (to measure  $|0_z\rangle$ ) and the  $s$  arm detector at time  $t_2$  (to measure  $|1_z\rangle$ ). We refer this variant as “ $xyz$ -six-state”.

## 6.2 The quantum space of the interferometric protocols

We assume Alice to be almost ideal, having the realistic space  $H^A = H_3$  (a qubit or a vacuum state), using two time-bin modes. As we have seen, four ancillary modes in vacuum states are added to each transmission. Therefore, the interferometer setups  $\mathcal{U}_{B_x}$  and  $\mathcal{U}_{B_y}$  transform the 2-mode states of  $H^A$  into a subspace that resides in the 6 modes space  $H^B$ . For simplicity, we assume that Eve does not generate  $n$ -photon states, with  $n \geq 2$ , so we can ignore high photon numbers in the  $H^B$  space<sup>12</sup>. Therefore, we redefine  $H^B = H_7$ , the space spanned by the vacuum, and the six single-photon terms in each of the above modes.

Using the  $x$  and  $y$  bases, Bob measures only time-bin  $t_1$ , so his actual measured spaces consist of two modes: time-bin  $t_1$  in the ‘ $s$ ’ arm and the ‘ $d$ ’ arm. In that case, the measured spaces are  $H^{B_x} = H^{B_y} = H_3$ , spanned by the states  $\{|000000\rangle_B^F, |010000\rangle_B^F, |000010\rangle_B^F\}$ . When Bob uses the  $z$  basis, he measures two different modes, so  $H^{B_z}$  is spanned by the states  $\{|000000\rangle_B^F, |000100\rangle_B^F, |001000\rangle_B^F\}$ .

Let us define the appropriate space  $H^{B^{-1}}$  for the 6-state protocol, according to Definition 8. The space  $H^{B^{-1}}$  is spanned by the states given by performing  $\mathcal{U} \in \{\mathcal{U}_{B_x}, \mathcal{U}_{B_y}\}$  on  $\{|000000\rangle_B^F, |010000\rangle_B^F, |000010\rangle_B^F\}$ , as well as the states given by performing  $\mathcal{U}_{B_z}$  on  $\{|000000\rangle_B^F, |000100\rangle_B^F, |001000\rangle_B^F\}$ . Interestingly, once applying  $\mathcal{U}^{-1}$ , the resulting states are embedded in an 8-mode space defined by the two incoming arms of the interferometer, ‘ $a$ ’ (from Alice) and ‘ $b$ ’ (from Bob), at time bins  $t'_{-1}$ ,  $t'_0$ ,  $t'_1$ , and  $t'_2$ . The basis states of  $H^{B^{-1}}$  are listed in Appendix B.3.

Following Definition 9, the QSoP  $H^P$  of this implementation for the 6-state protocol, is the subsystem of  $H^{B^{-1}}$  which is *controlled* by Eve. It is spanned by the 8-mode states spanning  $H^{B^{-1}}$  after tracing out Bob. The space that contains those “traced-out” states has only four modes that are controlled by Eve, specifically, input ‘ $a$ ’ of the interferometer at times  $t'_{-1}$  to  $t'_2$ , having a basis state of the form  $|a_{t'_{-1}} a_{t'_0} a_{t'_1} a_{t'_2}\rangle_P^F$ . Given the single-photon restriction, we get  $H^P = H_5$ , namely, the space spanned by the vacuum state, and a single photon in each of the four modes, i.e.  $\{|0000\rangle_P^F, |1000\rangle_P^F, |0100\rangle_P^F, |0010\rangle_P^F, |0001\rangle_P^F\}$ . This same result is obtained also if Bob measures all the six modes in  $H^B$ .

Bob might want to see how the basis states of the 4-mode QSoP,  $H^P$ , evolve through the interferometer in order to place detectors on the resulting modes, which will be used to identify Eve’s attack. It is interesting to note, that those basis states result in *10 different non-empty modes (!)*. If Bob measures all these modes, he *increases* the QSoP, and maybe allows Eve to attack a larger space, and so on and so forth. Therefore, in order to perform a security analysis, one must first fix the scheme and only then assess the QSoP. Otherwise,

<sup>12</sup>As mentioned in Section 2, this assumption is not legitimate when proving unconditional security of a protocol.

a “ping-pong” effect might increase the spaces’ dimensions to infinity. A similar, yet reversed logic, hints that it could actually be better for Bob, in terms of the simplicity of the analysis for the “*xy*-BB84” scheme, to measure *just* the two modes at  $t_1$  (i.e. the space spanned by  $|0, n_{s1}, 0, 0, n_{d1}, 0\rangle_B^F$ ), thus reducing the QSoP to a 2-mode space,  $H^P = H^A$ , see Appendix B.4. Although Eve is allowed to attack a larger space than this two-mode  $H^P$ , she has no advantage in doing so: pulses that enter the interferometer on different modes (i.e. other time-bins than  $t'_0$  and  $t'_1$ ), never interfere with the output pulses of time-bin  $t_1$  measured by Bob. Therefore, state occupying different modes can not be distinguished from the states in which those modes are empty.

### 6.3 The “Reversed-Space” attack on interferometric protocols

Consider a BB84 variant in which Bob uses only the  $x$  and the  $z$  bases, using a single interferometer, where the  $z$ -basis measurement is performed according to the description in the last few lines of Section 6.1. We refer this variant as “*xz*-BB84”. The QSoP of this scheme,  $H^P$  is the space described above for the “*xyz*-six-state” protocol. The following attack

$$|0\rangle_E |0100\rangle_P \xrightarrow{\mathcal{U}_E} \frac{1}{2} |E_0\rangle_E (|1000\rangle_P^F + |0100\rangle_P^F) + \frac{1}{2} |E_1\rangle_E (|0010\rangle_P^F + |0001\rangle_P^F) \quad (7)$$

$$|0\rangle_E |0010\rangle_P \xrightarrow{\mathcal{U}_E} \frac{1}{2} |E_1\rangle_E (-|1000\rangle_P^F + |0100\rangle_P^F) + \frac{1}{2} |E_0\rangle_E (|0010\rangle_P^F - |0001\rangle_P^F) \quad (8)$$

which we call “the Reversed-Space Attack”, allows Eve to acquire information about the transmitted qubits, without inducing *any* errors. The states  $|\cdot\rangle_E$  denote Eve’s ancilla which is not necessarily a photonic system. The state  $|0_z\rangle_A \equiv |0100\rangle_P^F$  and  $|1_z\rangle_A \equiv |0010\rangle_P^F$  are the regular states send by Alice, where we added the relevant extension of  $H^A$  in  $H^P$ . When  $|0_z\rangle_A$  is sent by Alice, the attacked state  $\mathcal{U}_E |0\rangle_E |0_z\rangle_A$  reaches Bob’s interferometer, and interferes in a way such that it can never reach Bob’s detector at time  $t_2$ , i.e.  ${}^F\langle 001000|_B \mathcal{U}_{B_x} ((\mathcal{U}_E |0\rangle_E |0_z\rangle_A) |0000\rangle_{B'}^F) = 0$ . Although the attacked state  $\mathcal{U}_E |0\rangle_E |0_z\rangle_A$  reaches modes that Alice’s original state  $|0_z\rangle_A$  can never reach, Bob never measures those modes, and cannot notice the attack. A similar argument applies when Alice sends  $|1_z\rangle_A$ .

As for the  $x$  basis<sup>13</sup>, this attack satisfies

$$|0\rangle_E |0_x\rangle_A \mapsto \frac{1}{\sqrt{8}} (|E_0\rangle_E + |E_1\rangle_E) (|0100\rangle_P^F + |0010\rangle_P^F) + \frac{1}{\sqrt{8}} (|E_0\rangle_E - |E_1\rangle_E) (|1000\rangle_P^F - |0001\rangle_P^F) \quad (9)$$

$$|0\rangle_E |1_x\rangle_A \mapsto \frac{1}{\sqrt{8}} (|E_0\rangle_E - |E_1\rangle_E) (|0100\rangle_P^F - |0010\rangle_P^F) + \frac{1}{\sqrt{8}} (|E_0\rangle_E + |E_1\rangle_E) (|1000\rangle_P^F + |0001\rangle_P^F). \quad (10)$$

The first element in the sum results in the desired interference in Bob’s lab, while the second is not measured by Bob’s detectors at time  $t_1$ . By letting Eve’s probes  $|E_0\rangle_E$  and  $|E_1\rangle_E$  be orthogonal states, Eve gets a lot of information while inducing no errors at all. Yet, we find that Eve is increasing the loss rate by this attack to 87.5%, but a very high loss rate is anyhow expected by Bob (as explained in the analysis of the PNS [13] and the tagging [24] attacks).

In conclusion, this attack demonstrates the risk of using various setups without giving full security analysis for the *specific* setup. We are not familiar with any other security analysis that takes into account the enlarged space generated by the inverse-transformation of Bob’s space.

<sup>13</sup>For simplicity we use the shorter notation  $|0_x\rangle \equiv (|0100\rangle_P^F + |0010\rangle_P^F)/\sqrt{2}$ , etc.

## 7 Conclusion

In this paper we have defined the QSA, a novel attack that generalizes all currently known attacks on the channel. This new attack brings a new method for performing security analysis of protocols. The attack is based on a realistic view of the quantum spaces involved, and in particular, the spaces that become larger than the theoretical ones, due to practical considerations. Although this paper is explicitly focused on the case of uni-directional implementations of a few schemes, its main observations and methods apply to any uni-directional QKD protocol, to bi-directional QKD protocols, and maybe also to any realistic quantum cryptography scheme beyond QKD.

The main conclusion of this research is that the quantum space which is attacked by Eve can be assessed, given a proper understanding of the experimental limitations. This assessment requires a novel cryptanalysis formalism — analyzing the states generated in Alice’s lab, as well as the states that are to be measured by Bob (assessing them as if they go backwards in time from Bob’s lab); this type of analysis resembles the two-time formalism in quantum theory [1, 44].

Open problems for further theoretical research include: 1.— Generalization of the QSA to other conventional protocols (such as the two-state protocol, EPR-based protocols, d-level protocols, etc.); such a generalization should be rather straightforward. 2.— Proving unconditional security (or more limited security results such as “robustness” [11]) against various QSAs. This is especially important for the interferometric setup, where the QSoP is much larger than Alice’s six-dimensional space (the one spanned by  $\chi^6$ ). 3.— Describing the QSA for more complex protocols, such as two-way protocols [33, 10, 11] in which the quantum communication is bi-directional, and protocols which use a larger set of states such as data-rejected protocols [2] or decoy-state protocols [25, 45, 30, 47]. 4.— Extend the analysis and results to composable QKD [4]. 5(a).— In some cases, if Bob uses “counters” and treats various measurement outcomes as errors, the effective QSoP relevant for proving security is potentially *much smaller* than the QSoP defined here. 5(b).— Adding counters on more modes increases the QSoP defined here, but might allow analysis of a smaller “attack’s QSoP”, if those counters are used to identify Eve’s attack. More generally, the connection between the way Bob interprets his measured outcomes, and the “attack’s QSoP” is yet to be further analyzed.

**Acknowledgments.** We thank Michel Boyer, Dan Kenigsberg and Hoi-Kwong Lo for helpful remarks.

## References

- [1] D. Z. Albert, Y. Aharonov, and S. D’Amato. Curious new statistical prediction of quantum mechanics. *Physical Review Letters*, 54(1):5–7, Jan 1985.
- [2] S. M. Barnett, B. Huttner, and S. J. D. Phoenix. Eavesdropping Strategies and Rejected-data Protocols in Quantum Cryptography. *Journal of Modern Optics*, 40:2501–2513, Dec. 1993.
- [3] H. Bechmann-Pasquinucci and N. Gisin. Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography. *Physical Review A*, 59(6):4238–4248, Jun. 1999.
- [4] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim. The universal composable security of quantum key distribution. In *TCC 2005: Second Theory of Cryptography Conference*, pages 386–406, Jan. 2005.
- [5] C. H. Bennett and G. Brassard. Quantum Cryptography: Public key distribution and coin tossing. *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, Dec. 1984.
- [6] E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. P. Roychowdhury. A proof of the security of quantum key distribution. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (STOC)*, pages 715–724, New York, 2000. ACM Press.
- [7] E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. P. Roychowdhury. A proof of the security of quantum key distribution. *J. Cryptology*, 19(4):381–439, 2006.



- [8] E. Biham, M. Boyer, G. Brassard, J. van de Graaf, and T. Mor. Security of Quantum Key Distribution Against All Collective Attacks. *Algorithmica*, 34:372–388, Nov. 2002.
- [9] E. Biham and T. Mor. Security of quantum cryptography against collective attacks. *Physical Review Letters*, 78(11):2256–2259, Mar 1997.
- [10] K. Boström and T. Felbinger. Deterministic secure direct communication using entanglement. *Physical Review Letters*, 89(18):187902, Oct 2002.
- [11] M. Boyer, D. Kenigsberg, and T. Mor. Quantum key distribution with classical Bob. ArXiv Quantum Physics e-prints, 2007. quant-ph/0703107.
- [12] G. Brassard, N. Lütkenhaus, T. Mor, and B. C. Sanders. Security Aspects of Practical Quantum Cryptography. In *EUROCRYPT 2000: International Conference on the Theory and Application of Cryptographic Techniques*, LNCS 1807:289–299, 2000.
- [13] G. Brassard, N. Lütkenhaus, T. Mor, and B. C. Sanders. Limitations on Practical Quantum Cryptography. *Physical Review Letters*, 85:1330–1333, Aug. 2000.
- [14] K. J. Blow, R. Loudon, S. Phoenix and T. J. Shepherd. Continuum fields in quantum optics *Physical Review A*, 42(7):4102–4114, Oct. 1990.
- [15] D. Bruß. Optimal Eavesdropping in Quantum Cryptography with Six States. *Physical Review Letters*, 81:3018–3021, Oct. 1998.
- [16] H. F. Chau. Practical scheme to share a secret key through a quantum channel with a 27.6% bit error rate. *Physical Review A*, 66(6):060302, Dec. 2002. For different (slightly smaller) numbers, see [24].
- [17] M. Dusek, N. Lutkenhaus, and M. Hendrych. Quantum Cryptography. ArXiv Quantum Physics e-prints, Jan. 2006. quant-ph/0601207.
- [18] C. Elliott, D. Pearson, and G. Troxel. Quantum cryptography in practice. In *SIGCOMM '03: Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 227–238, New York, NY, USA, 2003. ACM Press.
- [19] A. Ekert, B. Huttner, G. Palma and A. Peres. Eavesdropping on quantum-cryptographical systems. *Physical Review A*, 50(2):1047–1056, Aug. 1994.
- [20] C. Fuchs, N. Gisin, R. Griffiths, C.S. Niu and A. Peres. Optimal eavesdropping in quantum cryptography. I. Information bound and optimal strategy. *Physical Review A*, 56(2):1163–1172, Aug. 1997.
- [21] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy. Trojan-horse attacks on quantum-key-distribution systems. *Physical Review A*, 73(2):022320+, Feb. 2006.
- [22] N. Gisin, B. Kraus, and R. Renner. Lower and upper bounds on the secret key rate for QKD protocols using one-way classical communication. *Physical Review Letters*, 95:080501, 2005.
- [23] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. *Reviews of Modern Physics*, 74:145–195, Jan. 2002.
- [24] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill. Security of quantum key distribution with imperfect devices. *Quantum Information and Computation*, 5:325–360, 2004.
- [25] W.-Y. Hwang. Quantum key distribution with high loss: Toward global secure communication. *Physical Review Letters*, 91(5):057901, Aug. 2003.
- [26] W.-Y. Hwang, I.-T. Lim, and J.-W. Park. No-clicking event in quantum key distribution. ArXiv Quantum Physics e-prints, 2004. quant-ph/0412206.
- [27] H. Inamori, N. Lütkenhaus, and D. Mayers. Unconditional security of practical quantum key distribution. *European Physical Journal D*, 41:599–627, Mar. 2007.
- [28] H.-K. Lo and H. F. Chau. Unconditional security of quantum key distribution over arbitrarily long distances. *Science*, 283:2050–2056, 1999.
- [29] H.-K. Lo. Proof Of Unconditional Security of Six-State Quantum Key Distribution Scheme. *Quantum Information and Computation*, 1(2):81–94, Aug. 2001.
- [30] H.-K. Lo, X. Ma, and K. Chen. Decoy State Quantum Key Distribution. *Physical Review Letters*, 94(23):230504+, Jun. 2005.
- [31] V. Makarov, A. Anisimov, and J. Skaar. Effects of detector efficiency mismatch on security of quantum cryptosystems. *Physical Review A*, 74:022313, 2006.
- [32] C. Marand and P. Townsend Quantum key distribution over distances as long as 30 km *Optics Letters*, 20:1695–1697, Aug. 1995.
- [33] A. Muller, T. Herzog, B. Huttner, W. Tittel, H. Zbinden and N. Gisin. “Plug and play” systems for quantum cryptography. *Applied Physics Letters*, 70:793–395, Feb 1997.

- [34] V. Makarov and D. R. Hjelm. Faked states attack on quantum cryptosystems. *Journal of Modern Optics*, 52:691–705, May 2005.
- [35] D. Mayers. Unconditional security in quantum cryptography. *J. ACM*, 48(3):351–406, 2001, based on [46].
- [36] A. Niederberger, V. Scarani, and N. Gisin. Photon-number-splitting versus cloning attacks in practical implementations of the Bennett-Brassard 1984 protocol for quantum cryptography. *Physical Review A*, 71:042316, 2005.
- [37] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, UK, 2000.
- [38] A. Peres. How to differentiate between non-orthogonal states. *Physics Letters A*, 128:19, Mar 1988.
- [39] A. Peres. *Quantum Theory: concepts and methods*. Kluwer, Dordrecht, 1993.
- [40] V. Scarani, A. Acín, G. Ribordy and N. Gisin. Quantum Cryptography Robust against Photon Number Splitting Attacks for Weak Laser Pulse Implementations. *Physical Review Letters*, 92(5):057901–+, Feb. 2004.
- [41] M. Scully and M. S. Zubairy. *Quantum Optics*. Cambridge University Press, Cambridge, United Kingdom, 1997.
- [42] P. W. Shor and J. Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Physical Review Letters*, 85:441–444, 2000, based on [28].
- [43] P. D. Townsend. Secure key distribution system based on quantum cryptography. *Electronics Letters*, 30:809–811, May. 1994.
- [44] L. Vaidman, Y. Aharonov, and D. Z. Albert. How to ascertain the values of  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$  of a spin-1/2 particle. *Physical Review Letters*, 58(14):1385–1387, Apr. 1987.
- [45] X. B. Wang. Beating the Photon-Number-Splitting Attack in Practical Quantum Cryptography. *Physical Review Letters*, 94:230503, Jun. 2005.
- [46] A. Yao. Security of quantum protocols against coherent measurements *STOC '95: Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, pages 67–75, Las Vegas, Nevada, United States, 1995.
- [47] Z. L. Yuan, A. W. Sharpe, and A. J. Shields. Unconditionally secure one-way quantum key distribution using decoy pulses. *Applied Physics Letters*, 90:1118–+, Jan. 2007.
- [48] <http://www.idquantique.com/>
- [49] <http://www.maqtech.com/>

## Appendix

### A Mathematical Description of the PNS attack

The PNS attack can be realized using (an infinite set of) polarization independent beams splitters. Eve uses a beam splitter to split photons from Alice's state. Using a non-demolition measurement Eve measures the number of photons in one output of the beam splitter, and repeat the splitting until she acquires exactly one photon. Formally  $\mathcal{U}_E$  is defined:

$$\begin{aligned} |00\rangle_E^F |02\rangle_A^F &\mapsto |01\rangle_E^F |01\rangle_P^F & |00\rangle_E^F |10\rangle_A^F &\mapsto |10\rangle_E^F |00\rangle_P^F \\ |00\rangle_E^F |20\rangle_A^F &\mapsto |10\rangle_E^F |10\rangle_P^F & |00\rangle_E^F |01\rangle_A^F &\mapsto |01\rangle_E^F |00\rangle_P^F \\ |00\rangle_E^F |11\rangle_A^F &\mapsto (|01\rangle_E^F |10\rangle_P^F + |10\rangle_E^F |01\rangle_A^F) / \sqrt{2}. \end{aligned}$$

Whenever Alice sends a pulse with two photons of the same polarization, Eve and Bob end up, each, with having a single photon of the original polarization.

**Proposition 1.** *Eve's PNS attack for a pulse of 2 photons, gives Eve full information while inducing no errors.*

*Proof.* According to its definition it is trivial to verify the attack for the horizontal and vertical polarizations  $|0_z\rangle^{(2)}$  and  $|1_z\rangle^{(2)}$  (where  $|P\rangle^{(k)}$  means  $k$  photons having polarization  $P$ ). Using the standard creation and annihilation operators ( $a^\dagger$  and  $a$ )<sup>14</sup>, we can write the state of two photons in the diagonal polarization ( $x$  basis):  $|0_x\rangle^{(2)} = \left(\frac{1}{\sqrt{2}}(a_1^\dagger + a_2^\dagger)\right)^2 |00\rangle^F = \frac{1}{2}(|20\rangle^F + \sqrt{2}|11\rangle^F + |02\rangle^F)$ , similarly  $|1_x\rangle^{(2)} = \frac{1}{2}(|20\rangle^F - \sqrt{2}|11\rangle^F + |02\rangle^F)$ .

$$\begin{aligned} |00\rangle_E^F |0_x\rangle_P^{(2)} &\equiv \frac{1}{2} |00\rangle_E^F (|20\rangle^F + \sqrt{2}|11\rangle^F + |02\rangle^F)_P \\ &\xrightarrow{\mathcal{U}_E} \frac{1}{2} (|10\rangle_E^F |10\rangle_P^F + |01\rangle_E^F |10\rangle_P^F + |10\rangle_E^F |01\rangle_P^F + |01\rangle_E^F |01\rangle_P^F) \\ &= \frac{1}{2} ((|10\rangle_E^F + |01\rangle_E^F) |10\rangle_P^F + (|10\rangle_E^F + |01\rangle_E^F) |01\rangle_P^F) \\ &= \frac{1}{2} (|10\rangle_E^F + |01\rangle_E^F) (|10\rangle_P^F + |01\rangle_P^F) \\ &\equiv |0_x\rangle_E |0_x\rangle_P^{(1)} \\ \\ |00\rangle_E^F |1_x\rangle_P^{(2)} &\equiv \frac{1}{2} |00\rangle_E^F (|20\rangle^F - \sqrt{2}|11\rangle^F + |02\rangle^F)_P \\ &\xrightarrow{\mathcal{U}_E} \frac{1}{2} (|10\rangle_E^F |10\rangle_P^F - |01\rangle_E^F |10\rangle_P^F - |10\rangle_E^F |01\rangle_P^F + |01\rangle_E^F |01\rangle_P^F) \\ &= \frac{1}{2} ((|10\rangle_E^F - |01\rangle_E^F) |10\rangle_P^F - (|10\rangle_E^F - |01\rangle_E^F) |01\rangle_P^F) \\ &= \frac{1}{2} (|10\rangle_E^F - |01\rangle_E^F) (|10\rangle_P^F - |01\rangle_P^F) \\ &\equiv |1_x\rangle_E |1_x\rangle_P^{(1)} \end{aligned}$$

Which completes the proof. □

---

<sup>14</sup> See any quantum optics book, e.g. [41]

## A.1 Polarization change

A Polarization based QKD protocol makes a use of a Pockel cell ( $\mathcal{U}_{B_x}$ ), rotating the polarization of the photons going through it. For a single photon, its action is trivial,

$$\begin{aligned} |10\rangle^F &\xrightarrow{\mathcal{U}_{B_x}} \frac{1}{\sqrt{2}} (|10\rangle^F + |01\rangle^F), \text{ and} \\ |01\rangle^F &\xrightarrow{\mathcal{U}_{B_x}} \frac{1}{\sqrt{2}} (|10\rangle^F - |01\rangle^F). \end{aligned} \quad (11)$$

For a state that contains multiple photons, the transformation is not intuitive, and most simply defined using the creation and annihilation operators. In a somewhat simplified way, the Pockel cell can be considered as performing  $a_1^\dagger \mapsto \left(\frac{1}{\sqrt{2}}(a_1^\dagger + a_2^\dagger)\right)$  and  $a_2^\dagger \mapsto \left(\frac{1}{\sqrt{2}}(a_1^\dagger - a_2^\dagger)\right)$ , so that a state is transformed in the following way

$$|nm\rangle^F = \left(a_1^\dagger\right)^n \left(a_2^\dagger\right)^m |00\rangle^F \xrightarrow{\mathcal{U}_{B_x}} \left(\frac{1}{\sqrt{2}}(a_1^\dagger + a_2^\dagger)\right)^n \left(\frac{1}{\sqrt{2}}(a_1^\dagger - a_2^\dagger)\right)^m |00\rangle^F. \quad (12)$$

## B QSoP of the Interferometric Scheme: Supplementary Information

### B.1 A (brief) graphical description of pulses evolution through interferometer

See Figure 2 for evolution of a single occupied mode through the interferometer, and Figure 3 for evolution of two superpositioned modes.

### B.2 Evolution of modes through the interferometer

In order to simplify the analysis (a simplification that is not allowed when proving the full security of a scheme) we look at the ideal case in which exactly one photon (or none) is sent by Alice. The basis states are then the vacuum  $|000000\rangle_B^F \equiv |V\rangle_B^F$ , and the six states (that we denote for simplicity by)  $|100000\rangle_B^F \equiv |s_0\rangle_B^F$ ;  $|010000\rangle_B^F \equiv |s_1\rangle_B^F$ ;  $|001000\rangle_B^F \equiv |s_2\rangle_B^F$ ;  $|000100\rangle_B^F \equiv |d_0\rangle_B^F$ ;  $|000010\rangle_B^F \equiv |d_1\rangle_B^F$  and  $|000001\rangle_B^F \equiv |d_2\rangle_B^F$ .

The full transformation of a single photon pulse through the interferometer is given by Equation (5). Alice sends photons at time bins  $t'_0$  and  $t'_1$  only, so the interferometer transformation on Alice's basis states is  $|00\rangle_A^F |0000\rangle_{\hat{B}}^F \mapsto |V\rangle_B^F$ , and

$$\begin{aligned} |10\rangle_A^F |0000\rangle_{\hat{B}}^F &\mapsto (|s_0\rangle_B^F - e^{i\phi}|s_1\rangle_B^F + i|d_0\rangle_B^F + ie^{i\phi}|d_1\rangle_B^F) / 2 \\ |01\rangle_A^F |0000\rangle_{\hat{B}}^F &\mapsto (|s_1\rangle_B^F - e^{i\phi}|s_2\rangle_B^F + i|d_1\rangle_B^F + ie^{i\phi}|d_2\rangle_B^F) / 2, \end{aligned} \quad (13)$$

where  $|0000\rangle_{\hat{B}}$  denotes ancilla added during the process<sup>15</sup>. Equation 13 can be used to describe the interferometer effect on a general qubit, shown in Equation (6).

The states sent by Alice during the “xy-BB84” protocol evolve in the interferometer as follows:

$$\begin{aligned} |0_x\rangle_A &\xrightarrow{\phi=0} (|s_0\rangle_B^F - |s_2\rangle_B^F + i|d_0\rangle_B^F + 2i|d_1\rangle_B^F + i|d_2\rangle_B^F) / \sqrt{8} \\ |1_x\rangle_A &\xrightarrow{\phi=0} (|s_0\rangle_B^F - 2|s_1\rangle_B^F + |s_2\rangle_B^F + i|d_0\rangle_B^F - i|d_2\rangle_B^F) / \sqrt{8} \\ |0_y\rangle_A &\xrightarrow{\phi=\pi/2} (|s_0\rangle_B^F + |s_2\rangle_B^F + i|d_0\rangle_B^F - 2|d_1\rangle_B^F - i|d_2\rangle_B^F) / \sqrt{8} \\ |1_y\rangle_A &\xrightarrow{\phi=\pi/2} (|s_0\rangle_B^F - 2i|s_1\rangle_B^F - |s_2\rangle_B^F + i|d_0\rangle_B^F + i|d_2\rangle_B^F) / \sqrt{8} \end{aligned} \quad (14)$$

<sup>15</sup>Those ancillas (the space  $H^{\hat{B}}$ ) are originated by Alice extended space  $H^P$  and by Bob ( $H^{B'}$ ). Performing  $\mathcal{U}^{-1}$  reveals the exact origin of those ancillas.

Bob can distinguish the computation basis elements of bases  $x$  and  $y$ , measuring time-bin  $t_1$ , i.e. the states  $|d_1\rangle^F$  for  $|0\rangle$  and  $|s_1\rangle^F$  for  $|1\rangle$  in the measured basis. Other states give Bob no information about the state sent by Alice.

### B.3 $H^{B^{-1}}$ of the “ $xyz$ -six-state” scheme

Let Bob be using interferometric setups  $\mathcal{U}_{B_x}$  and measuring 6 modes (corresponding the space with a basis state  $|n_{s_0}n_{s_1}n_{s_2}n_{d_0}n_{d_1}n_{d_2}\rangle_B^F$ ) with one or less photons. Following Definition 8, the states spanning the space  $H^{B^{-1}}$  can be derived using Equation (6) (adjusted to the appropriate space):

$$\begin{aligned}
|000000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_x}^{-1}} |00000000\rangle_{PB'}^F \\
|010000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_x}^{-1}} \frac{1}{2} (-|01000000\rangle_{PB'}^F + |00100000\rangle_{PB'}^F - i|00000100\rangle_{PB'}^F - i|00000010\rangle_{PB'}^F) \\
|000010\rangle_B^F &\xrightarrow{\mathcal{U}_{B_x}^{-1}} \frac{1}{2} (-i|01000000\rangle_{PB'}^F - i|00100000\rangle_{PB'}^F + |00000100\rangle_{PB'}^F - |00000010\rangle_{PB'}^F) \\
|000000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_z}^{-1}} |00000000\rangle_{PB'}^F \\
|001000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_z}^{-1}} \frac{1}{2} (-|00100000\rangle_{PB'}^F + |00010000\rangle_{PB'}^F - i|00000010\rangle_{PB'}^F - i|00000001\rangle_{PB'}^F) \\
|000100\rangle_B^F &\xrightarrow{\mathcal{U}_{B_z}^{-1}} \frac{1}{2} (-i|10000000\rangle_{PB'}^F - i|01000000\rangle_{PB'}^F + |00001000\rangle_{PB'}^F - |00000100\rangle_{PB'}^F) \\
|000000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_y}^{-1}} |00000000\rangle_{PB'}^F \\
|010000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_y}^{-1}} \frac{1}{2} (i|01000000\rangle_{PB'}^F + |00100000\rangle_{PB'}^F - |00000100\rangle_{PB'}^F - i|00000010\rangle_{PB'}^F) \\
|000010\rangle_B^F &\xrightarrow{\mathcal{U}_{B_y}^{-1}} \frac{1}{2} (-|01000000\rangle_{PB'}^F - i|00100000\rangle_{PB'}^F - i|00000100\rangle_{PB'}^F - |00000010\rangle_{PB'}^F) \quad (15)
\end{aligned}$$

defined over the space  $H^P \otimes H^{B'}$  with basis state  $|a_{t'_{-1}}a_{t'_0}a_{t'_1}a_{t'_2}b_{t'_{-1}}b_{t'_0}b_{t'_1}b_{t'_2}\rangle_{PB'}^F$ . Note that performing  $\mathcal{U}^{-1}$  requires an additional ancilla, since the modes number increases from six to eight.

### B.4 QSoP of the “ $xy$ -BB84” scheme

Assume Bob measures only time-bin  $t_1$  in both output arms of the interferometer, i.e. the measured space is  $H^B$  subspace spanned by  $|0, n_{s_1}, 0, 0, n_{d_1}, 0\rangle_B^F$ . Assuming a single-photon restriction, the reversed space,

of that measured space that is spanned by:

$$\begin{aligned}
|000000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_x}^{-1}} |0000\rangle_{PB'}^F \\
|010000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_x}^{-1}} \frac{1}{2} (-|1000\rangle_{PB'}^F + |0100\rangle_{PB'}^F - i|0010\rangle_{PB'}^F - i|0001\rangle_{PB'}^F) \\
|000010\rangle_B^F &\xrightarrow{\mathcal{U}_{B_x}^{-1}} \frac{1}{2} (-i|1000\rangle_{PB'}^F - i|0100\rangle_{PB'}^F + |0010\rangle_{PB'}^F - |0001\rangle_{PB'}^F) \\
|000000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_y}^{-1}} |0000\rangle_{PB'}^F \\
|010000\rangle_B^F &\xrightarrow{\mathcal{U}_{B_y}^{-1}} \frac{1}{2} (i|1000\rangle_{PB'}^F + |0100\rangle_{PB'}^F - |0010\rangle_{PB'}^F - i|0001\rangle_{PB'}^F) \\
|000010\rangle_B^F &\xrightarrow{\mathcal{U}_{B_y}^{-1}} \frac{1}{2} (-|1000\rangle_{PB'}^F - i|0100\rangle_{PB'}^F - i|0010\rangle_{PB'}^F - |0001\rangle_{PB'}^F)
\end{aligned} \tag{16}$$

as can be verified using Equation (6). The space  $H^{B^{-1}}$  is embedded in a 4-mode space  $H^P \otimes H^{B'}$ , having the basis element  $|a_{t'_0} a_{t'_1} b_{t'_0} b_{t'_1}\rangle_{PB'}^F$ , i.e. Alice modes at times  $t'_0$  and  $t'_1$  and Bob's added ancillary modes at times  $t'_0$  and  $t'_1$  respectively. The resulting six states (16) span a 4-dimensional space, i.e.  $H^{B^{-1}} = H_4$ . The QSoP in this special case is  $H^P = H_3$ , spanned by  $|a_{t'_0} a_{t'_1}\rangle^F$  with one or less photons.

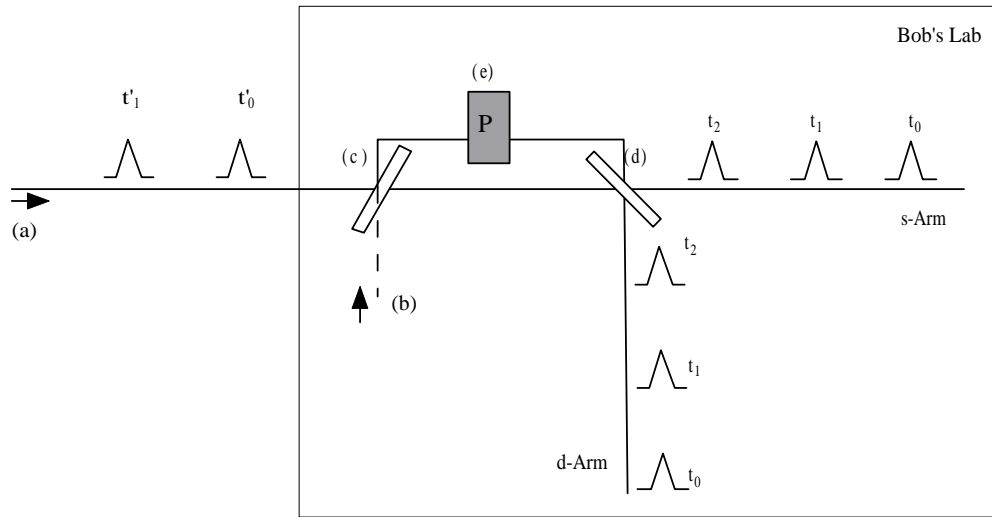
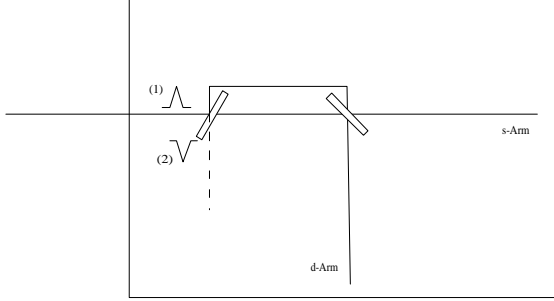
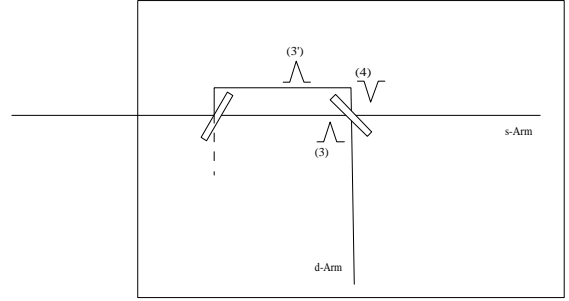


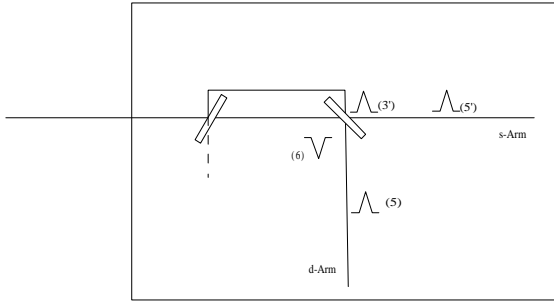
Figure 1: Bob's laboratory setup for the  $x$  and  $y$  basis. (a) Alice sends a qubit; (b) Vacuum states are added in the interferometer; (c), (d) beam-splitters; (e) phase shifter  $P_\phi$ .



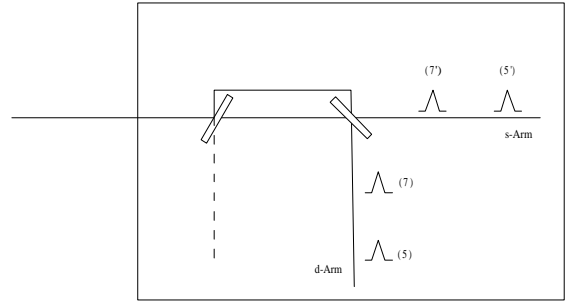
(a) Time  $T_0$ : the pulse (1) is about to enter the interferometer. A vacuum ancilla (2) is added in the input of the first beam splitter.



(b) Time  $T_1$ : Pulses (1) and (2) interfere and become a superposition of (3) and (3') in the short and long arms of the interferometer, respectively,  $|1\rangle_1|0\rangle_2 \xrightarrow{\text{BS}} (|1\rangle_3|0\rangle_{3'} + i|0\rangle_3|1\rangle_{3'})/\sqrt{2}$ . Pulse (3) is about to enter the second beam splitter so a vacuum ancilla is added (4).



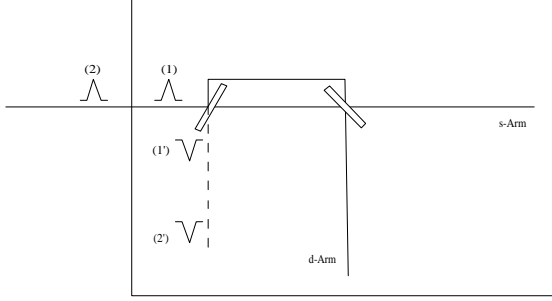
(c) Time  $T_2$ : pulses (5) and (5') are created by pulses (3) and (4),  $\frac{1}{\sqrt{2}}|0\rangle_4|1\rangle_3 \xrightarrow{\text{BS}} (i|1\rangle_5|0\rangle_{5'} + |0\rangle_5|1\rangle_{5'})/2$ . Pulse (3') is about to enter the second beam-splitter so a vacuum ancilla is added (6).



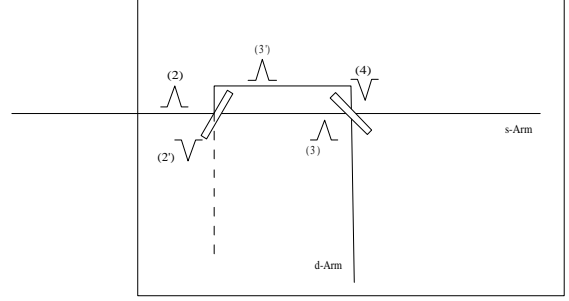
(d) Time  $T_3$ : Pulses (7) and (7') are created by interfering (3') and (6).  $\frac{i}{\sqrt{2}}|1\rangle_{3'}|0\rangle_6 \xrightarrow{\text{BS}} (i|1\rangle_7|0\rangle_{7'} - |0\rangle_7|1\rangle_{7'})/2$ .

Figure 2: Evolution in time of a single photon pulse through an interferometer satisfying  $|1000\rangle_{1,2,4,6} \xrightarrow{\text{Interferometer}} (|1000\rangle_{5',7',5,7} - |0100\rangle_{5',7',5,7} + i|0010\rangle_{5',7',5,7} + i|0001\rangle_{5',7',5,7})/2$ . The numbers represent the appropriate mode number of each pulse. The input state  $(|1\rangle_{t_0}|000\rangle)$  consists of modes (1) for the pulse at  $t_0$  and (2), (4) and (6) for the vacuum ancillas. The output modes that correspond to the state  $|n_{s_0}, n_{s_1}, n_{d_0}, n_{d_1}\rangle$  are modes (5'), (7'), (5) and (7) respectively.

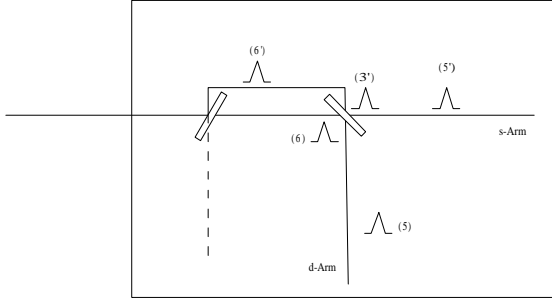




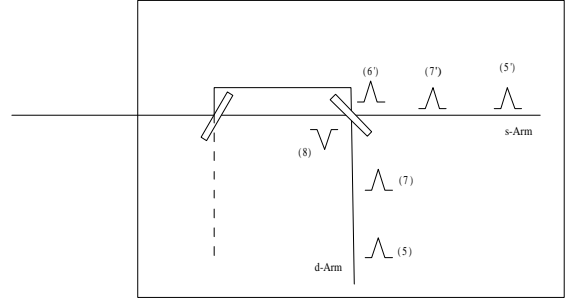
(a) Time  $T_0$ : The general single-photon qubit  $(\alpha|0\rangle + \beta|1\rangle)$  is sent to Bob is in two modes (1) and (2). Bob adds two vacuum ancillas (1') and (2') that interfere with the photon in the first beam splitter (BS-1).



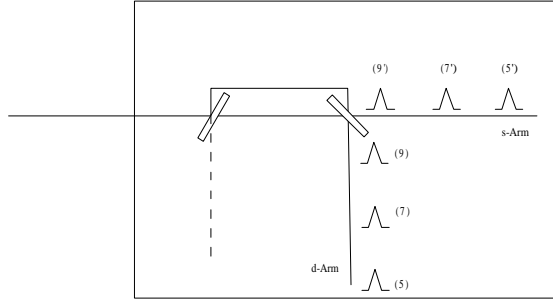
(b) Time  $T_1$ : Modes (1) and (1') interfere and create (3) and (3') in the short and long arm respectively,  $\alpha|1\rangle_1|0\rangle_{1'} \xrightarrow{\text{BS}} \frac{\alpha}{\sqrt{2}}(|1\rangle_3|0\rangle_{3'} + i|0\rangle_3|1\rangle_{3'})$ . Pulse (3) is about to enter BS-2 so a vacuum ancilla is added (4).



(c) Time  $T_2$ : Pulses (5) and (5') are created by the interference of (3) and (4)  $\frac{\alpha}{\sqrt{2}}|0\rangle_4|1\rangle_3 \xrightarrow{\text{BS}} \frac{i\alpha}{2}|1\rangle_5|0\rangle_{5'} + \frac{\alpha}{2}|0\rangle_5|1\rangle_{5'}$ . Pulses (6) and (6') created by the interference of (2) and (2') in BS-1  $\beta|1\rangle_2|0\rangle_{2'} \xrightarrow{\text{BS}} \frac{\beta}{\sqrt{2}}(|1\rangle_6|0\rangle_{6'} + i|0\rangle_6|1\rangle_{6'})$ .



(d) Time  $T_3$ : Pulses (7) and (7') are created by the interference of (3') and (6) in BS-2  $\frac{i\alpha}{\sqrt{2}}|1\rangle_{3'}|0\rangle_6 + \frac{\beta}{\sqrt{2}}|0\rangle_{3'}|1\rangle_6 \xrightarrow{\text{BS}} \frac{i(\alpha+\beta)}{2}|1\rangle_7|0\rangle_{7'} + \frac{\beta-\alpha}{2}|0\rangle_7|1\rangle_{7'}$ . Pulse (6') is about to enter BS-2 so a vacuum ancilla is added (8).



(e) Time  $T_4$ : Pulses (9) and (9') are created by the interference of (6') and (8) in BS-2  $\frac{i\beta}{\sqrt{2}}|1\rangle_{6'}|0\rangle_8 \xrightarrow{\text{BS}} \frac{i\beta}{2}|1\rangle_9|0\rangle_{9'} - \frac{\beta}{2}|0\rangle_9|1\rangle_{9'}$ .

Figure 3: Evolution in time of two modes through an interferometer satisfying  $(\alpha|1\rangle_1|0\rangle_2 + \beta|0\rangle_1|1\rangle_2)|0000\rangle_{1',2',4,8} \xrightarrow{\text{Interferometer}} (\frac{\alpha}{2}|100000\rangle + \frac{\beta-\alpha}{2}|010000\rangle - \frac{\beta}{2}|001000\rangle + \frac{i\alpha}{2}|000100\rangle + \frac{i(\alpha+\beta)}{2}|000010\rangle + \frac{i\beta}{2}|000001\rangle)_{5',7',9',5,7,9}$ . The numbers represent the appropriate mode number of each pulse. The corresponding state is  $|n_{s_0}n_{s_1}n_{s_2}n_{d_0}n_{d_1}n_{d_2}\rangle$ .